The Komlós Conjecture Holds for Vector Colorings

Aleksandar Nikolov*

January 18, 2013

Abstract

The Komlós conjecture in discrepancy theory states that for some constant K and for any $m \times n$ matrix $\mathbf{A}$ whose columns lie in the unit ball there exists a vector $\mathbf{x} \in \{-1, +1\}^n$ such that $\|\mathbf{Ax}\|_\infty \leq K$. This conjecture also implies the Beck-Fiala conjecture on the discrepancy of bounded degree hypergraphs. Here we prove a natural relaxation of the Komlós conjecture: if the columns of $\mathbf{A}$ are assigned unit vectors in $\mathbb{R}^n$ rather than ± 1 then the Komlós conjecture holds with $K = 1$. Our result rules out the possibility of a counterexample to the conjecture based on semidefinite programming. It also opens the way to proving tighter efficient (polynomial-time computable) upper bounds for the conjecture using semidefinite programming techniques.

1 Introduction

Let $\mathcal{H} = \{H_1, \ldots, H_m\}$ be a hypergraph with vertex set $V = [n]$. In this work we study the *combinatorial discrepancy* of hypergraphs and related quantities. The discrepancy of $\mathcal{H}$ is defined as

$$\mathrm{disc}(\mathcal{H}) = \min_{\chi:[n]\to\{-1,+1\}} \max_{i=1}^{m} \left| \sum_{j \in H_i} \chi(j) \right|.$$

Intuitively, discrepancy is the optimization problem of coloring the vertices of a hypergraph, so that the most imbalanced edge is as balanced as possible. Thus discrepancy is intimately connected to problems in Ramsey theory that study conditions under which every coloring leaves some edge monochromatic. Discrepancy has applications in geometry, computer science, and numerical integration, among others — the books by Matoušek [10], Chazelle [6], and the chapter by Beck and Sós [5] provide references for a wide array of applications.

We will be particularly interested in the discrepancy of hypergraphs with maximum degree bounded above by a parameter t, i.e. hypergraphs $\mathcal{H}$ all of

*Rutgers University, Piscataway NJ, email:anikolov@cs.rutgers.edu

whose vertices appear in at most t edges. It is a classical result of Beck and Fiala [4] that for any $\mathcal{H}$ of maximum degree at most t, $\mathrm{disc}(\mathcal{H}) \leq 2t-1$. Furthermore, they conjectured that $\mathrm{disc}(\mathcal{H}) \leq C\sqrt{t}$ for an absolute constant C. Proving Beck and Fiala's conjecture remains an elusive open problem in discrepancy theory.

As usual, we define the incidence matrix of $\mathcal{H}$ as an $m \times n$ 0-1 matrix $\mathbf{A}$ such that $A_{ij} = 1$ if and only if $j \in H_i$. In matrix notation discrepancy can be defined as $\mathrm{disc}(\mathcal{H}) = \min_{\mathbf{x} \in \{-1,1\}^n} \|\mathbf{A}\mathbf{x}\|_\infty$. This algebraic formulation allows us to extend the definition of discrepancy to arbitrary matrices: $\mathrm{disc}(\mathbf{A}) = \min_{\mathbf{x} \in \{-1,1\}^n} \|\mathbf{A}\mathbf{x}\|_\infty$. Interpreted in this way, discrepancy is a vector balancing problem: our goal is to assign signs to a given set of n vectors (the columns of $\mathbf{A}$), so that the signed sum has small norm (infinity norm in our case). A natural restriction on $\mathbf{A}$, analogous to the maximum degree restriction for hypergraphs, is to bound the maximum of some norm of the columns of $\mathbf{A}$. Such vector balancing problems were first considered in a general form by Bárány and Grinberg [3], although a similar problem was posed as early as 1963 by Dworetzky. The proof of Beck and Fiala shows that for any $\mathbf{A}$ whose columns have ℓ_1 norm at most 1, $\mathrm{disc}(\mathbf{A}) \leq 2$. Komlós conjectured[1] that for $\mathbf{A}$ whose columns have ℓ_2 norm at most 1, $\mathrm{disc}(\mathbf{A}) \leq K$ for some absolute constant K. The Komlós conjecture implies the Beck-Fiala conjecture and also remains open. The best partial progress towards proving the Komlós conjecture is a result by Banaszczyk [1], who showed the bound $\mathrm{disc}(\mathbf{A}) \leq K\sqrt{\log n}$ for an absolute constant K. This is the best known bound for the Beck-Fiala conjecture as well.

In this paper we are concerned with a natural convex relaxation of discrepancy: *vector discrepancy*. Vector discrepancy is defined analogously to discrepancy, but we "color" $[n]$ with unit n-dimensional vectors rather than ± 1:

$$\mathrm{vecdisc}(\mathbf{A}) = \min_{\mathbf{u_1},\ldots,\mathbf{u_n} \in S^{n-1}} \max_{i=1}^{m} \left\| \sum_{j=1}^{n} A_{ij}\mathbf{u_j} \right\|_2,$$

where S^{n-1} is the unit sphere in $\mathbb{R}^n$. Vector discrepancy is a relaxation of discrepancy, i.e. $\mathrm{vecdisc}(\mathbf{A}) \leq \mathrm{disc}(\mathbf{A})$ for all matrices $\mathbf{A}$: a coloring $\mathbf{x}$ achieving $\mathrm{disc}(\mathbf{A})$ induces a vector coloring $\{\mathbf{u_i} = x_i\mathbf{e_i}\}_{i=1}^n$ ($\mathbf{e_i}$ being the i-th standard basis vector) achieving vector discrepancy with the same value. Vector discrepancy was used by Lovász to give an alternative proof of Roth's lower bound on the discrepancy of arithmetic progressions [7]. A natural question is whether a lower bound on vector discrepancy could disprove the Komlós conjecture. Our main result is a negative answer to this question.

Theorem 1. *For any $m \times n$ real matrix $\mathbf{A}$ whose columns have ℓ_2 norm at most* 1, $\mathrm{vecdisc}(\mathbf{A}) \leq 1$.

[1] The earliest reference we can find is the 1987 book 'Ten Lectures on the Probabilistic Method' by Spencer [12]

Except as a means to lower bound discrepancy, vector discrepancy has also recently proved itself useful in establishing *efficient* upper bounds on discrepancy. In a recent breakthrough, Bansal [2] showed the following theorem.

Theorem 2 ([2]). *Let* $\mathbf{A}$ *be a real* $m \times n$ *matrix and assume that for any submatrix* $\mathbf{B}$ *of* $\mathbf{A}$ *we have* $\text{vecdisc}(\mathbf{B}) \leq D$. *Then* $\text{disc}(\mathbf{A}) \leq D \cdot K \log m$, *and, furthermore, there exists a polynomial time randomized algorithm which on input* $\mathbf{A}$ *outputs* $\mathbf{x} \in \{-1, 1\}^n$ *such that, with high probability,* $\|\mathbf{A}\mathbf{x}\|_\infty \leq D \cdot K \log m$ *for an absolute constant* K. □

In light of Bansal's result, Theorem 1 implies that for any $\mathbf{A}$ whose columns lie in the unit ball $\text{disc}(\mathbf{A}) \leq K \log m$ and that a coloring $\mathbf{x}$ achieving this bound can be found in randomized polynomial time. Such an efficient upper bound for the Komlós conjecture was proved by Bansal [2], and later using different methods by Lovett and Meka [9]. However, Bansal's, and Lovett and Meka's upper bounds are based on the "partial coloring" method and a $\log n$ factor seems inherent to upper bounds for the Komlós conjecture derived using this method. On the other hand, Matoušek [11] conjectures that the $\log m$ factor in Theorem 2 can be improved to $\sqrt{\log m}$. If this conjecture holds, we would have an alternative, and *efficient* proof of Banaszczyk's upper bound. We note that Banaszczyk's proof does not obviously yield an efficient algorithm, and no polynomial time algorithm that matches his bound is currently known.

To the best of our knowledge, Theorem 1 establishes the first constant upper bound on the vector discrepancy of matrices with bounded column ℓ_2 norms and on the vector discrepancy of bounded degree hypergraphs. A weaker bound of $O(\sqrt{\log m})$ can be derived in a variety of ways: directly from Banaszczyk's upper bound; from the existence of constant discrepancy partial colorings for the Komlós conjecture; from Matoušek's recent upper bound [11] on vector discrepancy in terms of the determinant lower bound of Lovász, Spencer, and Vesztergombi [8]. Our bound is tight, as $\text{vecdisc}((1)) = 1$, for example.

Techniques. Our proof of Theorem 1 relies on a dual characterization of vector discrepancy, first used by Matoušek to show that the determinant lower bound on discrepancy is almost tight [11]. However, our result does not follow directly from Matoušek's techniques, which only imply a bound of $O(\sqrt{\log m})$. Vector discrepancy is equivalent to a semidefinite programming problem, and, using a variant of the Farkas lemma for semidefinite programming, we can can formulate a dual program which is feasible for a parameter D precisely when $\text{vecdisc}(\mathbf{A}) \geq D$. We assume that the dual program is feasible for $D = 1 + \epsilon$. Geometrically, this feasibility can be formulated as the existence of two ellipsoids E and F such that $F \subseteq E$ and the sum of squared axes lengths of E is at most a D factor larger than the sum of squares axes lengths of F. The containment $F \subseteq E$ and the fact that the columns of $\mathbf{A}$ lie inside the unit ball imply that the largest k-dimensional section of E has volume lowerbounded by the largest k-dimensional section of F, for all k. Therefore, the axes lengths of E multiplicatively majorize the axes lengths of F, and Schur convexity implies a contradiction to the assumed constraints on the axes lengths of E and F.

2 Preliminaries

In this section we introduce some basic notation and useful linear algebraic facts.

2.1 Notation

We use boldface to denote matrices: $\mathbf{A}$, $\mathbf{X}$. We denote the entry in the i-th row and j-the column of $\mathbf{A}$ as A_{ij}. We denote by range($\mathbf{A}$) the vector space spanned by the columns of $\mathbf{A}$, and by ker($\mathbf{A}$) the kernel (nullspace) of $\mathbf{A}$. We'll assume a generic matrix $\mathbf{A}$ has dimensions m by n. By $\|\cdot\|$ we denote the standard ℓ_2 norm.

For a real symmetric matrix $\mathbf{X}$, we use $\mathbf{X} \succeq 0$ to denote that $\mathbf{X}$ is positive semidefinite.

For a real m by n matrix $\mathbf{A}$, we define the *discrepancy* of $\mathbf{A}$ as

$$\mathrm{disc}(\mathbf{A}) = \min_{x \in \{-1,1\}^n} \|\mathbf{A}\mathbf{x}\|_\infty.$$

We define the *vector discrepancy* of $\mathbf{A}$ as

$$\mathrm{vecdisc}(\mathbf{A}) = \min_{\mathbf{u_1},\ldots,\mathbf{u_n} \in S^{n-1}} \max_{i=1}^{m} \left\| \sum_{j=1}^{n} A_{ij}\mathbf{u_j} \right\|_2,$$

where S^{n-1} is the $(n-1)$-dimensional unit sphere in $\mathbb{R}^n$. As noted earlier, $\mathrm{vecdisc}(\mathbf{A}) \leq \mathrm{disc}(\mathbf{A})$ for all $\mathbf{A}$.

2.2 Dual Characterization of Vector Discrepancy

For each matrix $\mathbf{A}$, vecdisc($\mathbf{A}$) is defined as the minimum value of a convex function over a convex set, i.e. as the value of a convex optimization problem. In particular, $\mathrm{vecdisc}(\mathbf{A})^2$ can be written as the optimal solution to the *semidefinite program*

$$\min D \tag{1}$$
$$\text{subject to} \tag{2}$$
$$\forall 1 \leq i \leq m : (\mathbf{A}\mathbf{X}\mathbf{A}^\mathbf{T})_{ii} \leq D \tag{3}$$
$$\forall 1 \leq i \leq n : X_{ii} = 1 \tag{4}$$
$$\mathbf{X} \succeq 0. \tag{5}$$

To see the equivalence, write the vectors $\mathbf{u_1}, \ldots, \mathbf{u_n}$ forming a vector coloring as the columns of the matrix $\mathbf{U}$ and set $\mathbf{X} = \mathbf{U}^\mathbf{T}\mathbf{U} \succeq 0$. Also, by the Cholesky decomposition of positive semidefinite matrices, any $\mathbf{X} \succeq 0$ can be written as $\mathbf{X} = \mathbf{U}^\mathbf{T}\mathbf{U}$ where the columns of $\mathbf{U}$ are unit vectors and therefore give a vector coloring.

Using strong duality for convex programming, we can derive the dual program to (1)–(5) and characterize $\text{vecdisc}(\mathbf{A})^2$ as the optimal (maximum) solution to this dual. A derivation of the dual appears in recent work by Matoušek [11]. Next we present the resulting characterization of vector discrepancy. For a detailed proof of Theorem 3, see [11].

Theorem 3 ([11])**.** *For any real $m \times n$ matrix $\mathbf{A}$, $\text{vecdisc}(\mathbf{A}) \geq D$ if and only there exists a distribution p over $[m]$ and a vector $\mathbf{w} \in \mathbb{R}^n$ satisfying*

$$\sum_j w_j \geq D^2, \tag{6}$$

such that for all $\mathbf{z} \in \mathbb{R}^n$

$$\mathbb{E}_{i\sim p}\Big(\sum_{j=1}^{n} A_{ij} z_j\Big)^2 \geq \sum_{j=1}^{n} w_j z_j^2. \tag{7}$$

□

We note a geometric interpretation of Theorem 3. Define the ellipsoids $E(p, \mathbf{A}) = \{\mathbf{z} : \mathbb{E}_{i\sim p}(\sum_{j=1}^n A_{ij} z_j)^2 \leq 1\}$ and $F(\mathbf{w}) = \{\mathbf{z} : \sum_{j=1}^n w_j z_j^2 \leq 1\}$. Then Theorem 3 states that $\text{vecdisc}(\mathbf{A}) \geq D$ if and only if there exists a distribution p and $\mathbf{w}$ satisfying (6) such that $F(\mathbf{w}) \subseteq E(p, \mathbf{A})$.

2.3 Linear Algebra

The following two lemmas are essential to our proof. We suspect they are standard, but include detailed proofs for completeness. The first lemma states, geometrically, that any k-dimensional section of an ellipsoid E has volume upper bounded by the volume of the section with the subspace spanned by the k longest axes of E. We state and prove the lemma using elementary linear algebra.

Lemma 1. *Let $\mathbf{X} \in \mathbb{R}^{n\times n} : \mathbf{X} \succeq 0$ be a symmetric real matrix with eigenvalues $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$. Let also $\mathbf{U} \in \mathbb{R}^{n\times k}$ be a matrix with mutually orthogonal unit columns. Then $\det(\mathbf{U^T X U}) \leq \sigma_1 \ldots \sigma_k$.*

Proof. Let $\mathbf{U_1}$ and $\mathbf{U_2}$ be two matrices with mutually orthogonal unit columns such that $\text{range}(\mathbf{U_1}) = \text{range}(\mathbf{U_2})$. We claim that $\det(\mathbf{U_1^T X U_1}) = \det(\mathbf{U_2^T X U_2})$, and, therefore $\det(\mathbf{U^T X U})$ is entirely determined by $\text{range}(\mathbf{U})$. Indeed, there exists a unitary matrix $\mathbf{S}$ such that $\mathbf{U_1 S} = \mathbf{U_2}$, and, therefore, $\det(\mathbf{U_2^T X U_2}) = \det(\mathbf{S^T U_1^T X U_1 S}) = \det(\mathbf{U_1^T X U_1}) \det(\mathbf{S})^2 = \det(\mathbf{U_1^T X U_1})$.

The lemma is trivially true if $\det(\mathbf{U^T X U}) = 0$, so we will assume that $\mathbf{U^T X U}$ is non-singular. The proof of the lemma proceeds by induction on k. In the base case $k = 1$, the matrix $\mathbf{U}$ is just a unit vector $\mathbf{u}$. By the min-max characterization of eigenvalues we have that for any unit vector $\mathbf{u}$, $\det(\mathbf{u^T X u}) = \mathbf{u^T X u} \leq \sigma_1$. Let us assume the lemma holds for $k-1$. Let $\mathbf{v_1}, \ldots, \mathbf{v_n}$ be the eigenvectors of $\mathbf{X}$ associated with $\sigma_1, \ldots, \sigma_n$. Since $\text{range}(\mathbf{U})$ is a vector space of dimension k, we have that $\text{range}(\mathbf{U}) \cap \text{span}\{\mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a vector space

of dimension at least $k-1$. Let $\mathbf{U_0} \in \mathbb{R}^{n \times k-1}$ be a matrix whose columns are an orthonormal basis for some $k-1$ dimensional vector space contained in $\mathrm{range}(\mathbf{U}) \cap \mathrm{span}\{\mathbf{v_2}, \ldots, \mathbf{v_n}\}$. Let $\mathbf{P}$ be a projection matrix for the space $\mathrm{span}\{\mathbf{v_2}, \ldots, \mathbf{v_n}\}$. Since the columns of $\mathbf{U_0}$ are elements of $\mathrm{span}\{\mathbf{v_2}, \ldots, \mathbf{v_n}\}$, we have that $\mathbf{PU_0} = \mathbf{U_0}$, and, therefore, $\mathbf{U_0^T P^T X P U_0} = \mathbf{U_0^T X U^0}$. Also, the top $n-1$ eigenvalues of $\mathbf{P^T X P}$ are $\sigma_2, \ldots, \sigma_n$, and, by the induction hypothesis, $\det(\mathbf{U_0^T X U_0}) \leq \sigma_2 \ldots \sigma_n$. Let $\mathbf{u}$ be a unit vector in $\mathrm{range}(\mathbf{U_0})^\perp \cap \mathrm{range}(\mathbf{U})$ and let $\mathbf{U_1}$ be the matrix resulting from appending $\mathbf{u}$ as a column of $\mathbf{U_0}$. Since $\mathrm{range}(\mathbf{U}) = \mathrm{range}(\mathbf{U_1})$, we have that $\det(\mathbf{U^T X U}) = \det(\mathbf{U_1^T X U_1})$. A direct calculation shows that

$$\mathbf{U_1^T X U_1} = \begin{pmatrix} \mathbf{U_0^T X U_0} & \mathbf{U_0^T X u} \\ \mathbf{u^T X U_0} & \mathbf{u^T X u} \end{pmatrix}$$

Therefore,

$$\det(\mathbf{U_1^T X U_1}) = (\mathbf{u^T X u}) \det(\mathbf{U_0^T X U_0} - \frac{\mathbf{U_0^T X u u^T X U_0}}{\mathbf{u^T X u}}).$$

Let $\mathbf{v} = \frac{1}{\sqrt{\mathbf{u^T X u}}} \mathbf{U_0^T X u}$. Since $\mathbf{X}$ is symmetric, the equality above reduces to $\det(\mathbf{U_1^T X U_1}) = (\mathbf{u^T X u}) \det(\mathbf{U_0^T X U_0} - \mathbf{v^T v})$. Since we assumed that $\mathbf{U^T X U}$ is non-singular, we know that $\mathrm{range}(\mathbf{U_0}) \cap \ker(\mathbf{X}) \subseteq \mathrm{range}(\mathbf{U}) \cap \ker(\mathbf{X}) = \emptyset$, and, therefore, $\mathbf{U_0^T X U_0}$ is also non-singular. By the matrix determinant lemma, $\det(\mathbf{U_0^T X U_0} - \mathbf{v^T v}) = \det(\mathbf{U_0^T X U_0})(1 - \mathbf{v^T}(\mathbf{U_0^T X U_0})^{\mathbf{-1}}\mathbf{v})$. Since $\mathbf{U_0^T X U_0} \succeq 0$, $(\mathbf{U_0^T X U_0})^{-1} \succeq 0$, and therefore $\mathbf{v^T}(\mathbf{U_0^T X U_0})^{\mathbf{-1}}\mathbf{v} \geq 0$. So we have that $\det(\mathbf{U_1^T X U_1}) \leq (\mathbf{u^T X u}) \det(\mathbf{U_0^T X U_0})$. By the min-max characterization of eigenvalues, $\mathbf{u^T X u} \leq \sigma_1$ and this completes the inductive step. ☐

Lemma 2. *Let* $\mathbf{X} \in \mathbb{R}^{n \times n} : \mathbf{X} \succeq 0$ *and* $\mathbf{Y} \in \mathbb{R}^{n \times n} : \mathbf{Y} \succeq 0$. *Suppose that*

$$\forall \mathbf{u} \in \mathbb{R}^n : \mathbf{u^T X u} \geq \mathbf{u^T Y u}.$$

Then, $\det(\mathbf{X}) \geq \det(\mathbf{Y})$.

Proof. For a symmetric real matrix $\mathbf{M} \succeq 0$, define the ellipsoid $E(\mathbf{M}) = \{\mathbf{u} : \mathbf{u^T M u} \leq 1\}$. $E(\mathbf{M})$ is unbounded if and only if $\mathbf{M}$ is singular. Otherwise,

$$\mathrm{vol}(E(\mathbf{M})) = \frac{\mathrm{vol}(B^n)}{\sqrt{\det(\mathbf{M})}}, \tag{8}$$

where B^n is the n-dimensional unit ball.

By assumption, $E(\mathbf{X}) \subseteq E(\mathbf{Y})$. If $\det(\mathbf{Y}) = 0$, the lemma is trivially true. If $\det(\mathbf{X}) = 0$, then $E(\mathbf{X})$ is unbounded and therefore $E(\mathbf{Y})$ is unbounded, which implies $\det(\mathbf{Y}) = 0$. If, on the other hand, $E(\mathbf{X})$ and $E(\mathbf{Y})$ are bounded, we have that $\mathrm{vol}(E(\mathbf{X})) \leq \mathrm{vol}(E(\mathbf{Y}))$, and, by (8), $\det(\mathbf{X}) \geq \det(\mathbf{Y})$, as desired. ☐

3 Proof of Main Theorem

We begin with an inequality which can be seen as a converse to the geometric mean–arithmetic mean inequality. The inequality follows from the Schur convexity of symmetric convex functions; we present a self-contained elementary proof using a powering trick.

Lemma 3. *Let $x_1 \geq \ldots \geq x_n > 0$ and $y_1 \geq \ldots \geq y_n > 0$ such that*

$$\forall k \leq n : x_1 \ldots x_k \geq y_1 \ldots y_k \tag{9}$$

Then,

$$\forall k \leq n : x_1 + \ldots + x_k \geq y_1 + \ldots + y_k. \tag{10}$$

Proof. We will show that for all positive integers L, $(x_1 + \ldots + x_n)^L \geq \frac{1}{n!}(y_1 + \ldots + y_n)^L$. Taking L-th roots, we get that $x_1 + \ldots + x_n \geq \frac{1}{(n!)^{1/L}}(y_1 + \ldots + y_n)$. Letting $L \to \infty$ and taking limits yields the desired result.

By the multinomial theorem,

$$(x_1 + \ldots + x_n)^L = \sum_{i_1 + \ldots + i_n = L} \frac{L!}{i_1! \ldots i_n!} x_1^{i_1} \ldots x_n^{i_n}.$$

The inequalities (9) imply that whenever $i_1 \geq \ldots \geq i_n$, $x_1^{i_1} \ldots x_n^{i_n} \geq y_1^{i_1} \ldots y_n^{i_n}$. Therefore,

$$(x_1 + \ldots + x_n)^L \geq \sum_{\substack{i_1 \geq \ldots \geq i_n \\ i_1 + \ldots + i_n = L}} \frac{L!}{i_1! \ldots i_n!} y_1^{i_1} \ldots y_n^{i_n}. \tag{11}$$

Given a sequence $i_1, \ldots, i_n$, let σ be a permutation on n elements such that $i_{\sigma(1)} \geq \ldots \geq i_{\sigma(n)}$. Since $y_1 \geq \ldots \geq y_n$, we have that $y_1^{i_{\sigma(1)}} \ldots y_n^{i_{\sigma(n)}} \geq y_1^{i_1} \ldots y_n^{i_n}$. Furthermore, there are at most $n!$ distinct permutations of $i_1, \ldots, i_n$ (the bound is achieved exactly when all $i_1, \ldots, i_n$ are distinct). These observations and the multinomial theorem imply that

$$(y_1 + \ldots + y_n)^L \leq \sum_{\substack{i_1 \geq \ldots \geq i_n \\ i_1 + \ldots + i_n = L}} \frac{n! L!}{i_1! \ldots i_n!} y_1^{i_1} \ldots y_n^{i_n}. \tag{12}$$

Inequalities (11) and (12) together imply $(x_1 + \ldots + x_n)^L \geq \frac{1}{n!}(y_1 + \ldots + y_n)^L$ as desired. ∎

We are now ready to prove our main result.

Theorem 4 (Theorem 1 restated)**.** *For any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ such that $\forall i \in [n] : \|\mathbf{A}_{*\mathbf{i}}\| \leq 1$, $\mathrm{vecdisc}(\mathbf{A}) \leq 1$.*

Proof. We will use Theorem 3 with $D = \sqrt{1+\epsilon}$ for an arbitrary $\epsilon > 0$. For any $\mathbf{w} \in \mathbb{R}^n$ satisfying $\sum_{i=1}^n w_i \geq 1 + \epsilon$ we will show there exists a $\mathbf{z} \in \mathbb{R}^n$ satisfying

$$E_{j \sim p}(\sum_{i=1}^n A_{ij} z_i)^2 < \sum_i w_i z_i^2. \tag{13}$$

Therefore, by Theorem 3, $\mathrm{vecdisc}(A)^2 < 1+\epsilon$ for all $\epsilon > 0$, which proves our main theorem.

For any $i : w_i \le 0$, we can set $z_i = 0$. Then $\sum_{i:w_i>0} w_i \ge \sum_i w_i \ge 1+\epsilon$. Consider also the submatrix $\mathbf{A'}$ consisting of those columns $\mathbf{A_{*i}}$ of $\mathbf{A}$ for which $w_i \ge 0$. The matrix $\mathbf{A'}$ satisfies the assumption that all its columns have norm bounded by 1. Therefore, it is sufficient to show that for any matrix $\mathbf{A}$ with columns bounded by 1 in the euclidean norm, any $\mathbf{w}$ such that $\forall i : w_i > 0$ and $\sum_{i=1}^n w_i \ge 1+\epsilon$, and any distribution p on $[m]$, there exists a $\mathbf{z}$ satisfying (13).

We denote by $\mathbf{W}$ the diagonal matrix with $\mathbf{w}$ on the diagonal, and similarly for any distribution $\mathbf{p} \in \mathbb{R}^m_+ : \sum_{j=1}^m p_j = 1$ we denote by $\mathbf{P}$ the diagonal matrix with $\mathbf{p}$ on the diagonal. In this matrix notation, we need to show that for any positive definite diagonal matrix $\mathbf{W}$ such that $\mathrm{Tr}(\mathbf{W}) \ge 1+\epsilon$, and any positive semidefinite diagonal matrix $\mathbf{P}$ such that $\mathrm{Tr}(\mathbf{P}) = 1$, there exists a vector $\mathbf{z} \in \mathbb{R}^n$ such that $\mathbf{z^T A^T P A z} < \mathbf{zWz}$.

Assume for contradiction that

$$\forall \mathbf{z} : \mathbf{z^T A^T P A z} \ge \mathbf{zWz}. \tag{14}$$

Geometrically, this is equivalent to $F(p,\mathbf{A}) \subseteq E(\mathbf{w})$, where F and E are defined as before. The outline of our proof is as follows. The relation $F(p,\mathbf{A}) \subseteq E(\mathbf{w})$ implies that, for all k, the largest k-dimensional section of F has volume lower bounded by the volume of the largest k-dimensional section of E. Using Lemma 1 and the Hadamard bound we can show that this implies that, for all k, the product of the k largest p_i is lower bounded by the product of the k largest w_j. Then, Lemma 3 implies that the sum of all p_i is lower bounded by the sum of all w_j, which is a contradiction. We proceed to prove the above claims formally.

Let, without loss of generality, $w_1 \ge \ldots \ge w_n > 0$ and similarly $p_1 \ge \ldots \ge p_m \ge 0$. Denote by $\mathbf{A_{[k]}}$ the matrix $(\mathbf{A_{*1}}, \ldots, \mathbf{A_{*k}})$ and by $\mathbf{W_k}$ the diagonal matrix with $w_1, \ldots, w_k$ on the diagonal. We first show that

$$\forall k \le n : \det(\mathbf{A^T_{[k]} P A_{[k]}}) \le p_1 \ldots p_k. \tag{15}$$

Let $\mathbf{u_1}, \ldots \mathbf{u_k}$ be an orthonormal basis for the range of $\mathbf{A_{[k]}}$ and let $\mathbf{U_k}$ be the matrix $(\mathbf{u_1}, \ldots \mathbf{u_k})$. Then $\mathbf{A_{[k]}} = \mathbf{U_k U_k^T A_{[k]}}$. Each column of the square matrix $\mathbf{U_k^T A_{[k]}}$ has norm at most 1, and, by Hadamard's inequality, $\det(\mathbf{A^T_{[k]} U_k}) = \det(\mathbf{U_k^T A_{[k]}}) \le 1$. Therefore,

$$\forall k \le n : \det(\mathbf{A^T_{[k]} P A_{[k]}}) \le \det(\mathbf{U_k^T P U_k}).$$

By Lemma 1, we have that $\det(\mathbf{U_k^T P U_k}) \le p_1 \ldots p_k$, which proves (15).

By (14) we know that for all k and for all $\mathbf{u} \in \mathbb{R}^k$, $\mathbf{u^T A^T_{[k]} P A_{[k]} u} \ge \mathbf{u^T W_k u}$, since we can freely choose $\mathbf{z}$ such that $z_i = 0$ for all $i > k$. Then, by Lemma 2, we have that

$$\forall k \le n : \det(\mathbf{A^T_{[k]} P A_{[k]}}) \ge \det(\mathbf{W_k}) = w_1 \ldots w_k \tag{16}$$

Combining (15) and (16), we have that

$$\forall k \leq n : p_1 \dots p_k \geq w_1 \dots w_k \tag{17}$$

By Lemma 3, (17) implies that $1 = \sum_{j=1}^{m} p_j \geq \sum_{j=1}^{n} p_j \geq \sum_{i=1}^{n} w_i \geq 1 + \epsilon$, a contradiction.

□

4 Conclusion

We have shown that the vector discrepancy of a matrix **A** all of whose columns are contained in the unit ball is bounded by 1 from above. This result establishes a natural vector discrepancy variant of the notorious Komlós and Beck-Fiala conjectures. On one hand our result can be seen as evidence in support of the conjectures: they cannot be disproved by lower bounding vector discrepancy. On the other hand, our work opens the possibility of giving an efficient proof of Banaszczyk's bound of $O(\sqrt{\log m})$ on $\mathrm{disc}(\mathbf{A})$ by improving the pseudoapproximation algorithm of Bansal [2]. We hope that our result would prove useful in an attack on the Komlós conjecture itself.

Acknowledgements

I would like to thank Nikhil Bansal, Kunal Talwar, Daniel Dadush, and S. Muthukrishnan for useful discussions of the result and the writeup.

References

[1] W. Banaszczyk. Balancing vectors and gaussian measures of n-dimensional convex bodies. *Random Structures & Algorithms*, 12(4):351–360, 1998.

[2] N. Bansal. Constructive algorithms for discrepancy minimization. In *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, FOCS '10, pages 3–10, Washington, DC, USA, 2010. IEEE Computer Society.

[3] I. Bárány and VS Grinberg. On some combinatorial questions in finite-dimensional spaces. *Linear Algebra and its Applications*, 41:1–9, 1981.

[4] J. Beck and T. Fiala. Integer-making theorems. *Discrete Applied Mathematics*, 3(1):1–8, 1981.

[5] J. Beck and V.T. Sós. Discrepancy theory. In *Handbook of combinatorics (vol. 2)*, page 1446. MIT Press, 1996.

[6] B. Chazelle. *The Discrepancy Method.* Cambridge University Press, 1991.

[7] L. Lovász. Integer sequences and semidefinite programming. *Publ. Math. Debrecen*, 56:475–479, 2000.

[8] L. Lovász, J. Spencer, and K. Vesztergombi. Discrepancy of set-systems and matrices. *European Journal of Combinatorics*, 7(2):151–160, 1986.

[9] S. Lovett and R. Meka. Constructive discrepancy minimization by walking on the edges. *Foundations of Computer Science, IEEE Annual Symposium on*, 0:61–67, 2012.

[10] J. Matousek. *Geometric Discrepancy: An Illustrated Guide*. Springer Verlag, 2010.

[11] J. Matoušek. The determinant bound for discrepancy is almost tight. *Manuscript, Arxiv*, 1101:0767, 2011.

[12] J. Spencer. *Ten lectures on the probabilistic method*. SIAM, 1994.